Immunogenicity of Therapeutic Proteins

Therapeutic proteins enter the market in increasing numbers, and current observations confirm that most of these biopharmaceuticals are to a variable extent immunogenic. This may compromise the efficacy, safety as well as the target market size, and therefore this issue should not be neglected during drug development. We review the causes that can lead to immunogenicity and focus more specifically on the role that T-cell epitopes play in raising an immune response. An overview is given of the fundamental in vitro and in silico techniques that can be used to assess T-cell driven immunogenicity. Finally, we discuss how these techniques can be integrated into the drug development pipeline.

By Dr Ivo Van Walle, Dr Ignace Lasters and Philippe Stas

Immunogenicity is typically viewed as the presence of antibodies directed against the biotherapeutic upon administration of the therapeutic. These antibodies can affect the efficacy of the drug in multiple ways, including reducing the life time of the drug in the patient, neutralise its activity and alter its pharmacokinetics. Drug immunogenicity can lead to the cross-reactive antibodies that recognise both the drug and an autologous protein, leading to a depletion of the naturally occurring protein in a systemic way. Therefore, regulatory instances demand to document the putative immunogenicity of biotherapeutics during the drug development cycle and monitor immunogenicity throughout the clinical trials.

The causes of immunogenicity can be extrinsic or intrinsic to the protein. Extrinsic factors are drug formulation, aggregate formation, degradation products, contaminants and dosing. The administration mode, as well as the drug regime also strongly influence how immunogenicity should be assessed. That is, immunogenicity will have different effects for drugs that are given in acute indications compared to drugs to treat chronic diseases. In the latter case, patients are exposed to the drug over a longer period of time and as such can mount a complete response.

Intrinsic factors are related to the protein itself. These are mainly the T- and B-cell epitopes that are recognised specifically by the immune system, and which will be explained in the next section. An assessment of the intrinsic immunogenicity of a therapeutic protein is especially important when it has an endogenous counterpart. This may lead to the generation of cross-reactive antibodies against the endogenous counterpart, leading in turn to systemic depletion of the endogenous protein as well as the therapeutic protein. Also, fusion proteins created from fully human domains may give rise to unexpected immunogenicity, due to the introduction of epitopes in the linking region.

In recent years, whitepapers have been published on the design of assays to monitor the immunogenicity in vivo. While many of the mentioned aspects are difficult to model and predict beforehand, one specific driver of immunogenicity can be readily modelled and assessed early in the discovery cycle: the T-cell epitopes that are defined in the sequence of the biotherapeutic, and which are a necessary – though not sufficient – step in the mounting of an immune response.
**T-cell epitopes**

The immune system can react in two ways to proteins that it recognises as non-self. In a humoral response, antibodies are produced, and in a cellular response, cytotoxic T-cells (CTLs) will kill infected or cancerous cells that express the protein in question. Both types of response require the presence of one or more, resp, TH- or CTL epitopes in the protein. These are sequential regions of the protein (antigen) that, in the correct immunological context, will be recognised by resp a T-helper (TH) cell or a CTL. For biotherapeutics, it is mainly the humoral response that is important to investigate further, ie where the TH-epitopes are involved. The cytotoxic pathway and CTL epitopes are important only for intracellularly expressed proteins.

**TH-epitopes**

An antibody response involves three major types of cells: professional antigen presenting cells (APCs), TH-cells and B-cells. The APCs are typically dendritic cells, which will take up protein in a non-specific manner through endocytosis. In the endosome, this antigen is cleaved into a mixture of peptides with a length of up to 34 residues. The endosome then merges with a vesicle released from the Golgi apparatus, which contains membrane-bound Major Histocompatibility Complex (MHC) – in human these are called Human Leukocyte Antigen (HLA) – class II receptors (Figure 1a).

The HLAII receptor has the ability to bind a range of peptides, derived from the endocytosed protein, in an almost linearly outstretched fashion, and present it as such on the APC's surface after merger of the vesicle with the cell membrane (Figure 1b). However, each peptide has a different affinity for the receptor because of its particular amino acid sequence, and those that do not bind strongly enough will not be displayed on the cell surface.

Those peptides that do eventually bind to HLAII, the TH-epitopes, will be displayed on the APC's cell surface. This will continue until the epitope dissociates from the receptor, after which the latter becomes unstable and is internalised again. As such, the concentration of a particular epitope that is being presented, is a function of the amount loaded and, crucially, of its affinity for HLAII.

The HLAII/epitope complex on the surface may be bound by the T-cell receptor (TCR) of a TH-cell, a process which occurs inside the lymph nodes. The TCR is, like an antibody, unique to the cell in question, and therefore this is a first specific recognition of the antigen. The TH-cell population as a whole has been trained not to respond to peptides that are derived from autologous (self) proteins. A non-self peptide however can trigger the activation of a specific TH-cell, which will then lead to its proliferation and differentiation, such that it can stimulate those B-cells that produce antibodies against the protein as a whole.

**HLA diversity**

The presentation of TH- or CTL-epitopes on HLA is an absolute requirement for generating a specific immune response, except for a few T-independent antigens that will not be discussed further here. Antigens that do not contain epitopes, ie whose derived peptides are removed during processing or only have low affinity for HLA, will therefore escape. This potential weakness of the immune system is circumvented by the presence of several HLA receptor allotypes on the cell surface, each with different peptide-specificity.
The diversity is generated (i) by the presence of several HLA I and II genes, and (ii) by a very high degree of polymorphism of most of these genes on the level of the whole population. The differences between alleles are also almost exclusively located in the peptide binding site, illustrating the evolutionary benefit of having HLAs with different peptide affinities. Along the same line, specific HLAs have also been linked to diseases, such as rheumatoid arthritis, coeliac disease, insulin-dependent diabetes mellitus and narcolepsy. Therefore, assessment of the immunogenicity of a therapeutic protein or vaccine (lead) should take into account the prevalence of the different variants within the patient population that is being targeted.

Population coverage
The immunogenicity of a drug can be expressed as the percentage of the population that develops an immune response to it, for which we will here use the term coverage. For a vaccine, coverage should be as high as possible. The inverse applies to a therapeutic protein, where the protein should ideally not contain any TH-epitopes against any of the HLAII expressed by the individual in question. Coverage is, however, not straightforward to calculate, even when the actual epitopes are known. This is due to the fact that the frequencies of HLA allotypes (and haplotypes) can be quite different for different regions/ethnicities around the world, and that strong linkage exists between some alleles from different genes. To illustrate this, Table 1 gives an overview of frequencies of DR1 type receptors for different ethnicities. Note that the sum of all frequencies sums up to slightly less than 200% because both parents contribute.

Tests of immunogenicity
Peptides derived from antigens must bind to HLA with sufficient affinity, or there will be no specific immune response. Therefore, determining the affinity of these peptides, for those HLA allotypes that are (most) prevalent in the patient population, provides an overview of the expected immunogenicity. Such an overview is by definition a safe overestimation of the true immunogenicity, since not all epitopes are available after processing, or are able to stimulate T-cells.

In silico approaches
Peptides bind on HLA in an outstretched fashion. This greatly limits the number of possible binding modes of any peptide, as well as the interactions between its side-chains, up to a point where it becomes feasible to build a model that predicts affinity for a particular HLA allotype.

The earliest such models were statistical. Given a number of peptides with experimentally known affinity, the sequences can conceptually be aligned on the consecutive ‘positions’ in the HLA binding groove. Given the final alignment, a statistical profile, or matrix, is then constructed for the binding groove positions, and any new peptide can be aligned to that. The alignment score of the peptide against the matrix is then a measure of its affinity. Examples of such statistical methods are Rankpep, Propred, Tepitope, Bimas, Syphpeiti and others.

More elaborate models based solely on inference have been constructed as well, using neural networks and classification trees. Structure-based methods are more recent. These directly model the molecular interactions between peptide and receptor, using force-fields. As such, there is no need per se for experimental data on which to train the model. This has the advantage of avoiding to overfit the data of the model towards such experimental data, which is a serious problem in statistical methods. For example, a matrix with 10 positions has 10 x 20 = 200 parameters to fit, ie the score for each amino acid at each position. However, only a handful of the HLA allotypes have

### Table 1: Population frequencies of some highly prevalent HLA class II DR1 receptors in different ethnicities. Data compiled from 220 DR1 frequency studies for different regions/ethnicities

<table>
<thead>
<tr>
<th>HLAII DR</th>
<th>CAUCASIAN</th>
<th>ORIENTAL</th>
<th>BLACK</th>
</tr>
</thead>
<tbody>
<tr>
<td>DRB1*0101</td>
<td>14</td>
<td>523</td>
<td>4</td>
</tr>
<tr>
<td>DRB1*0301</td>
<td>17</td>
<td>8</td>
<td>12</td>
</tr>
<tr>
<td>DRB1*0701</td>
<td>24</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>DRB1*0803</td>
<td>1</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>DRB1*0901</td>
<td>3</td>
<td>23</td>
<td>5</td>
</tr>
<tr>
<td>DRB1*1101</td>
<td>9</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>DRB1*1102</td>
<td>1</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>DRB1*1202</td>
<td>1</td>
<td>20</td>
<td>0</td>
</tr>
<tr>
<td>DRB1*1301</td>
<td>10</td>
<td>3</td>
<td>13</td>
</tr>
<tr>
<td>DRB1*1302</td>
<td>7</td>
<td>3</td>
<td>14</td>
</tr>
<tr>
<td>DRB1*1404</td>
<td>0</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>DRB1*1501</td>
<td>17</td>
<td>19</td>
<td>4</td>
</tr>
<tr>
<td>DRB1*1503</td>
<td>1</td>
<td>0</td>
<td>18</td>
</tr>
</tbody>
</table>
more than 50 peptides of known affinity, and the total search space is 20^{10} peptides. Structure-based approaches such as Epibase circumvent this problem and can calculate peptide affinities for any HLA allotype, provided that a good model of its structure can be created. For most allotypes this is feasible given careful homology modelling.

The accuracy of predictive methods has evolved strongly over the past few years. In Figure 2 the higher accuracy of structure-based methodologies is illustrated by comparing three widely used methods, Rankpep, Propred and Epibase, on six of the major HLAII receptors.

Experimental approaches
Peptides binding to HLA are T-cell epitopes if the HLA-peptide complex can be productively recognised by the T-cell receptor expressed on CD4+ T-cells (recognition of TH-epitopes) or CD8+ T-cells (recognition of CTL epitopes). Focusing here on TH-epitopes, a number of assays have been developed to characterise the T-cell responsiveness. Typically peripheral blood mononuclear cells (PBMC) from patients or naïve community donors are harvested and primed with antigen. After a number of days of culture, the cells are restimulated with autologous PBMC and antigen or peptides. Typically, peptides derived from the biotherapeutic are tested in a 15-mer format, with subsequent peptides having a 12 residue overlap. The T-cell stimulation can then be determined by a suitable proliferation assay. Specific set-ups use APCs that are differentiated from blood monocytes. Subsequently, these cells are co-cultured with purified CD4+ T-cells and T-cell proliferation is assessed by thymidine incorporation. Routinely, T-cell activation is measured through ELISPOT analysis using an ELISA approach to measure the numbers of IFN-γ or IL-2 secreting T-cells. To obtain meaningful results it is mandatory to employ a large group of donors, typically more than 50, in order to cover the HLA allotype diversity to account for intrinsic variation in responsiveness across individuals, and to accurately measure low response levels for individual peptides. Such low response levels per peptide are expected in preclinical experiments as these are generally carried out with PBMC from naïve, community donors.

Examples of immunogenic therapeutic proteins
A rapidly growing number of therapeutic proteins are being used as therapeutic agents in the clinic. To reduce the risk of mounting an immune response, typically detected by the presence of circulating antibodies against the biopharmaceutical, it has generally been accepted that it should resemble as much as possible human, self proteins. However it has become clear that nearly all protein therapeutics, including fully human ones such as Epo, may be

Figure 2
Comparison of experimental (X-axis) versus predicted results (Y-axis) of Rankpep, Propred and Epibase, for six highly prevalent DR I receptors. Rankpep units are not shown, as they differ widely between receptors, and Propred values lower than -8 were set to this value. The affinity threshold for peptides that can actually be epitopes, is usually taken around -35kJ/mol (~800nM). Dataset compiled from patents WO 03/040165, WO 02/19986, and publicly available sources. The R^2 values are calculated only on peptides that can bind, taken here below -30kJ/mol (~800nM). Non-binding peptides are often given a high dummy value because of the impossibility to measure them accurately, explaining the vertical lines above -20kJ/mol.
Therapeutics

References

Immunogenic. This applies to human antibodies as well. For example, antibody A33, a fully humanised IgG1 mAb derived by CDR grafting elicits a human anti-human antibody (HAHA) response in eight of 11 patients. Apart from the presence of T\(_4\)-epitopes, the immunogenic response depends on a number of factors as reviewed by Koren et al, such as route of administration (subcutaneous versus intravenous injection), dosage regime and product aggregation. Importantly, the extent of humoral response raised against the biopharmaceutical can also depend on the context of usage. For example, patients with non-Hodgkin’s lymphoma do not develop HACA antibodies against Rituximab, a chimeric anti-CD20 monoclonal antibody but in patients with systemic lupus erythematosus, administration of Rituximab to deplete B-cells leads to a detectable HACA response in 11/17 patients, six of which develop high titer HACAs. Comparable data is available for Campath, where the observed immunogenicity depends on the indication treated. Figure 3 summarises a recent study of HAMA, HACA and HAHA levels of engineered antibodies.

Incorporating immunogenicity screening in the drug development pipeline
Ideally, immunogenicity screening should be performed at the lead selection stage. At that stage, the immunogenic risk of different leads can be used as one of several criteria to guide the selection. Also, this is still early enough in the drug development cycle to deimmunise the most promising leads, should they turn out to have a high immunogenic risk. On the other hand, immunoprofiling of biotherapeutics that have already entered the clinical development path allows an improved interpretation of the data obtained during the clinical trials. Moreover, it allows to document immunogenicity drivers such as reactive HLA allotypes.

The combination of in silico predictions followed by experimental confirmation is quite advantageous to determine the immunogenic risk, for two reasons. Firstly, costs can be lowered significantly by reducing the number of required – and expensive – experiments. Peptides that are very unlikely to be epitopes, as well as HLA allotypes (donors) that are very unlikely to react, can be omitted in this way. Secondly, experimental data do not have a very high degree of confidence by themselves, because of their low signal-to-noise level and other factors influencing the results, such as undersampling and the immunogenic background of the donors. Rationalisation, using a thorough knowledge of the HLA system and good predictive models is required, for example to trace back peptides to actual allotypes, and finally to arrive at a good estimate of the immunogenic risk.

Dr Ivo Van Walle is Senior Scientist at AlgoNomics. Prior to working with AlgoNomics, Dr Van Walle obtained his PhD in bioinformatics at the University of Brussels, where he specialised in bioinformatics algorithm design and multiple sequence and structure alignment techniques. Dr Van Walle holds an MScE in biotechnology.

Dr Ignace Lasters is Chief Executive Officer at AlgoNomics. Dr Lasters has more than 20 years’ industrial experience in protein engineering and structural bioinformatics. Since 1984, he was responsible for the protein design group of PGS (currently Bayer CropScience) which was acquired by Corvas International in 1991. Prior to setting up AlgoNomics in 1999, he worked as a consultant for the Genoscope subsidiary of Perkin Elmer. Dr Lasters received his PhD in chemistry at the Free University of Brussels.

Philippe Stas is Chief Operating Officer at AlgoNomics. Since 1990, Mr Stas has been active in antibody engineering projects, immunology and bioinformatics research at the University of Brussels. After working as IT-Manager for PE Genoscope (currently Applera, CA), he joined the team for the creation of AlgoNomics in 1999. Mr Stas holds an MSE in biotechnology and a Master degree in Information Technology.